**基于 AI 算法的医学图像辅助诊断**

1. **目的+实验方法概述**

### （一）目的

* 深入掌握深度学习算法在医学图像辅助诊断中的应用，尤其是与医学特征提取的结合。
* 熟练熟悉医学图像预处理与特征提取方法，理解其在医学图像分析中的重要性。
* 全面了解医学影像诊断的基本流程与标准，为医学特征提取提供应用背景支撑。
* 重点攻克深度学习算法选择与医学图像特征提取难题，提升特征提取的有效性。
* 全力克服医学图像数据集的构建与模型训练的难点，确保模型性能。

### （二）实验方法概述

1. 采用**迁移学习**策略，以在大规模图像数据集（ImageNet）上预训练的

**DenseNet121** 模型为基础，冻结初始层后构建自定义分类头；

1. 结合**数据增强**技术扩充训练数据多样性，缓解样本不平衡问题；
2. 通过**模型微调**（Fine-Tuning）进一步适配肺炎 X 光图像诊断任务；
3. 利用**交叉熵损失函数、类别权重调整、学习率调度、早停**等策略优化训练过程，最终通过测试集评估与单 / 多样本测试验证模型性能。
4. **环境**

### 2.1 硬件环境

* CPU：Intel Core i7
* GPU：NVIDIA GeForce RTX 3060（6G）
* 内存：16GB

### 2.2 软件环境

* 操作系统：Windows 11
* 编程语言：Python 3.8
* 深度学习框架：TensorFlow 2.11
* 其他库：OpenCV 用于图像处理，NumPy 用于数值计算，Matplotlib 用于数据可视化，PyQt5 用于开发肺炎诊断应用程序。

1. **基本原理**

**1、医学影像诊断基本流程与标准**

**1.1 基本流程**

1. **图像采集**：运用 X 光设备精准获取患者胸部的 X 光图像，图像质量直接影响后续诊断与特征提取。
2. **图像预处理**：对采集到的图像实施去噪、增强等处理，去除图像中的噪声干扰，提升图像清晰度与对比度，为特征提取奠定良好基础。
3. **特征提取**：从预处理后的图像中提取对肺炎诊断有价值的特征，如形状、纹理等特征，这些特征是区分正常与病变图像的关键。
4. **诊断决策**：依据提取的特征进行综合诊断，给出准确的诊断结果。

### 1.2 标准

肺炎诊断主要依据 X 光图像中肺部的阴影、纹理等特征。正常肺部图像纹理清晰、结构规则；肺炎患者肺部图像常出现片状阴影、纹理模糊、结构紊乱等表现，这些特征变化是特征提取重点关注的对象。

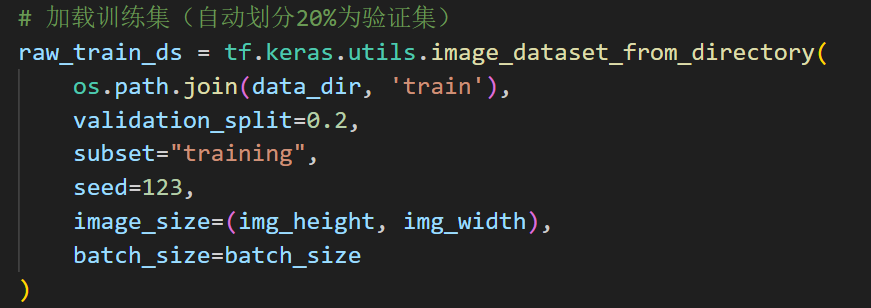
**2、数据集构建**

### 2.1 数据来源（数据集为肺炎患者X光片，数量约为1w张，格式为JPEG）

### 2.2 数据划分

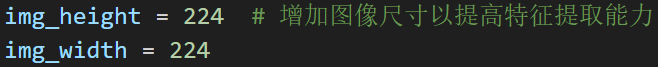
自动将数据集按 8:1:1 的比例划分为训练集、验证集和测试集。训练集用于模型参数学习，验证集用于调整模型超参数，测试集用于评估模型最终性能，确保模型在不同数据子集上都能得到有效验证。

训练集代码如下示（验证集、测试集基本相同，不再赘述）：



### 2.3 数据预处理

1. **调整尺寸**：将所有图像统一调整为 224x224 大小，满足深度学习模型输入要求，保证数据格式一致性。



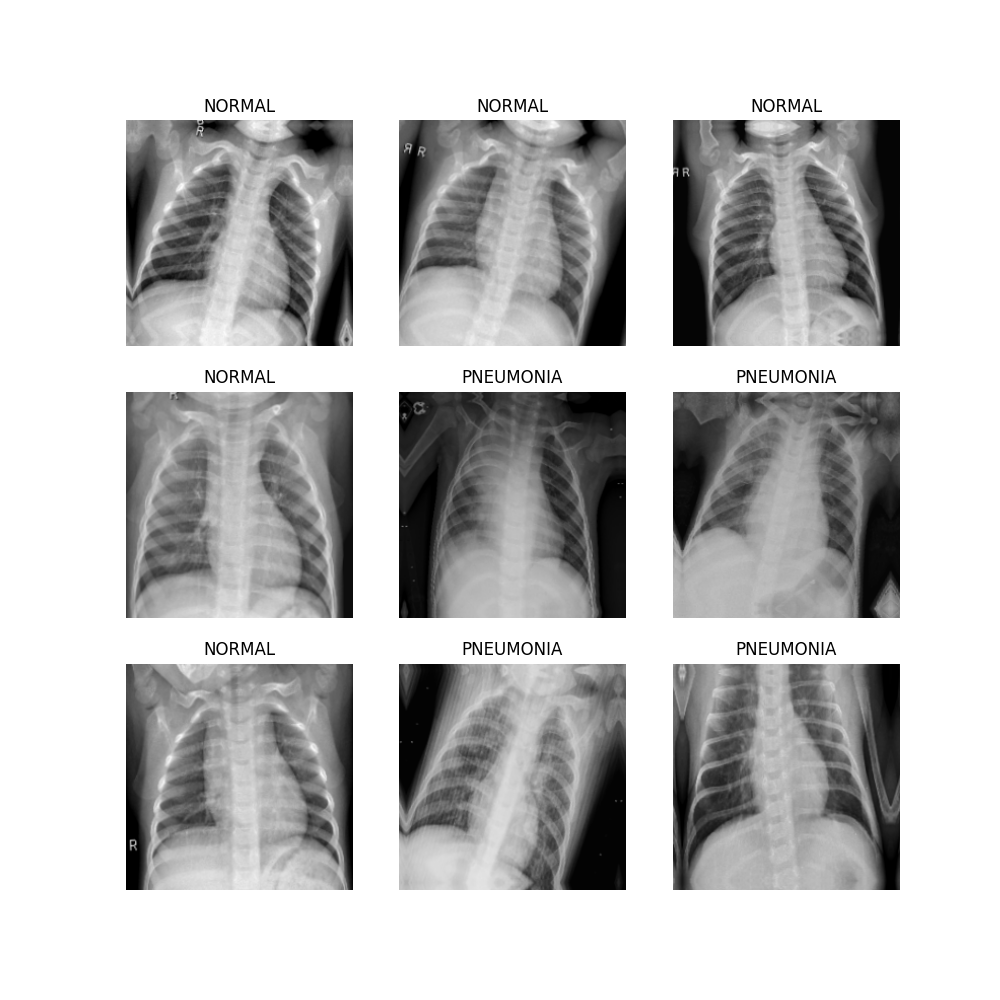
1. **数据增强**：对训练集图像进行随机翻转、旋转、缩放等操作，扩充数据多样性，提升模型泛化能力，使模型对不同角度、大小的图像特征都能有效学习。
2. **标准化**：将图像像素值从 [0, 255] 归一化到 [-1, 1]，消除图像亮度差异对特征提取的影响，让模型学习更聚焦于图像的本质特征。



在训练过程中会展示处理后的图片（仅展示前9张）



训练过程中的处理后的图像：



## 3、深度学习算法选择与模型构建

### 3.1 模型选择

选定 **DenseNet121** 作为基础模型，主要基于以下考量：

1. **预训练权重优势**：DenseNet121 在 ImageNet 数据集上经过预训练，已学习到大量通用图像特征，能快速适应医学图像特征提取任务，减少模型训练时间与数据需求。
2. **网络结构适配性**：其独特的**网络结构**可有效捕捉医学图像中的复杂特征，通过密集连接方式，加强特征传播与复用，提升特征提取效率。

### 3.2 模型构建

在 DenseNet121 基础上搭建模型，具体结构如下：

1. **标准化层**：将输入图像像素值归一化到 [-1, 1]，统一数据分布，利于模型训练收敛。
2. **DenseNet121**：充当特征提取器，前期冻结大部分层，仅训练最后几层，借助其预训练能力快速提取医学图像基础特征，后期逐步解冻更多层进行微调，优化特征提取效果。
3. **全局平均池化层**：压缩特征图空间维度为一维，降低数据维度，提取全局特征信息，减少计算量。
4. **批归一化层**：加速模型收敛，提升模型训练稳定性，减少内部协变量偏移，使模型训练更高效。
5. **Dropout 层**：随机丢弃部分神经元连接，防止模型过拟合，增强模型泛化能力，避免模型对训练数据过学习。
6. **全连接层**：完成分类预测任务，输出层采用 sigmoid 激活函数，适用于二分类场景，输出肺炎或正常的预测结果。

 DenseNet121 模型的加载、冻结策略及自定义分类头的构建代码截图：





## 4、医学图像特征提取

### 4.1 手工特征提取方法

### ****（4.1为理论分析，本次设计不考虑该种特征提取方法，本次设计的特征提取方法详见4.2）****

1. **形状特征提取**：理论上利用边缘检测算法，如 Canny 算子，提取肺部区域的边缘轮廓，计算轮廓周长、面积、圆形度等参数。例如，正常肺部轮廓相对规则，肺炎病变时可能导致肺部轮廓局部变形、面积变化等，这些形状特征变化可作为诊断依据。
2. **纹理特征提取**：理论上采用灰度共生矩阵（GLCM），计算图像中不同灰度值像素对在特定方向和距离上的出现频率，获取对比度、相关性、能量、熵等纹理特征。肺炎患者肺部纹理与正常情况相比，纹理复杂度、方向性等会发生改变，GLCM 提取的特征能有效反映这些变化。

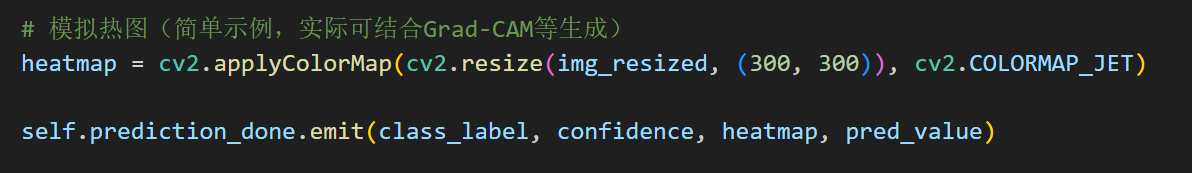
### 4.2 基于深度学习的特征提取

1. **DenseNet121 特征层次分析**：
   * **浅层特征（1 - 10 层）**：主要提取边缘、角点等基础特征，可识别肺部与周围组织边界、肋骨边缘等结构，帮助区分图像中的不同区域，为后续深入特征提取奠定基础。
   * **中层特征（11 - 50 层）**：学习到纹理基元、小区域模式等特征，能够捕捉肺部正常纹理结构以及早期病变可能引起的纹理细微变化，如肺纹理增粗、紊乱等特征，对早期肺炎诊断有重要意义。
   * **深层特征（51 - 120 层）**：聚焦于复杂语义特征，可识别肺炎典型病灶特征，如磨玻璃影、实变区域等，这些深层特征直接与肺炎病变相关，对准确诊断肺炎起到关键作用。

篇幅有限，仅展示模型部分构建过程



1. **特征可视化验证**：运用 Grad - CAM++ 技术生成类激活图，直观展示模型在图像上的关注区域。对于肺炎样本，激活区域多集中在病变部位，如肺部炎症浸润区域；正常样本激活区域均匀分布于全肺野。通过可视化验证，可评估特征提取效果，进一步优化模型对医学特征的学习。



## 5、模型训练

### 5.1 损失函数与优化器

1. **损失函数**：选用**二元交叉熵损失函数**，针对二分类任务，能有效衡量模型预测结果与真实标签之间的差异，引导模型优化方向。
2. **优化器**：采用 Adam 优化器，初始学习率设为 1e - 5，它结合了 Adagrad 和 RMSProp 算法优点，能自适应调整学习率，在训练过程中快速收敛到较优解。

### 5.2 训练策略

1. **分层训练**：先冻结 DenseNet121 所有层，单独训练顶层全连接层 20 个 epoch，让模型初步学习分类任务；之后解冻 DenseNet121 最后 20 层，以 1e - 6 较小学习率微调 10 个 epoch，使模型在已有预训练特征基础上，针对医学图像特征进一步优化，提升模型性能。
2. **回调函数**：运用 ReduceLROnPlateau 回调函数，根据验证集损失动态调整学习率，当损失在一定 epoch 内不再下降时，降低学习率，防止模型陷入局部最优；使用 EarlyStopping 回调函数，监控验证集指标（如准确率、召回率），当指标在一定 epoch 内无明显提升时，提前终止训练，避免过拟合；通过 ModelCheckpoint 回调函数，保存训练过程中的最佳模型，方便后续使用。

二元交叉熵损失函数”“Adam 优化器”“分层训练”“类别权重调整” 及回调函数策略的具体实现





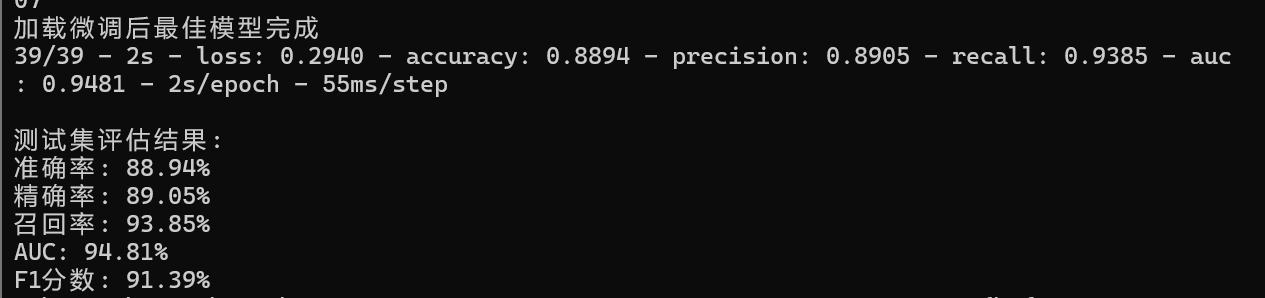
### 5.3 类别权重

鉴于数据集存在类别不平衡问题，正常样本数量较多，计算类别权重。增加正常样本权重，使模型在训练时更加关注少数类别的肺炎样本，平衡不同类别对模型训练的影响，提升模型对肺炎样本的识别能力。

## 6、实验结果与分析

### 6.1 模型评估指标

训练代码中采用**准确率**、**精确率**、**召回率**、**AUC** 和 **F1 分数**作为评估指标，具体结果如下：



这些指标综合反映模型在肺炎诊断任务中的性能，

准确率体现模型整体判断准确性---精确率衡量预测为肺炎样本中实际为肺炎的比例---召回率关注实际肺炎样本被正确预测的比例---AUC 评估模型对正负样本的区分能力---F1 分数综合精确率和召回率，更全面评估模型性能。

### 6.2 可视化训练过程 + 训练曲线分析

绘制训练过程中的准确率、损失、精确率和召回率曲线。从曲线可看出，模型训练初期，随着 epoch 增加，准确率逐步上升，损失逐渐下降，表明模型在不断学习；微调阶段，各指标进一步优化，验证了分层训练策略有效性。同时，通过曲线波动情况，可分析模型训练稳定性，及时调整训练参数。

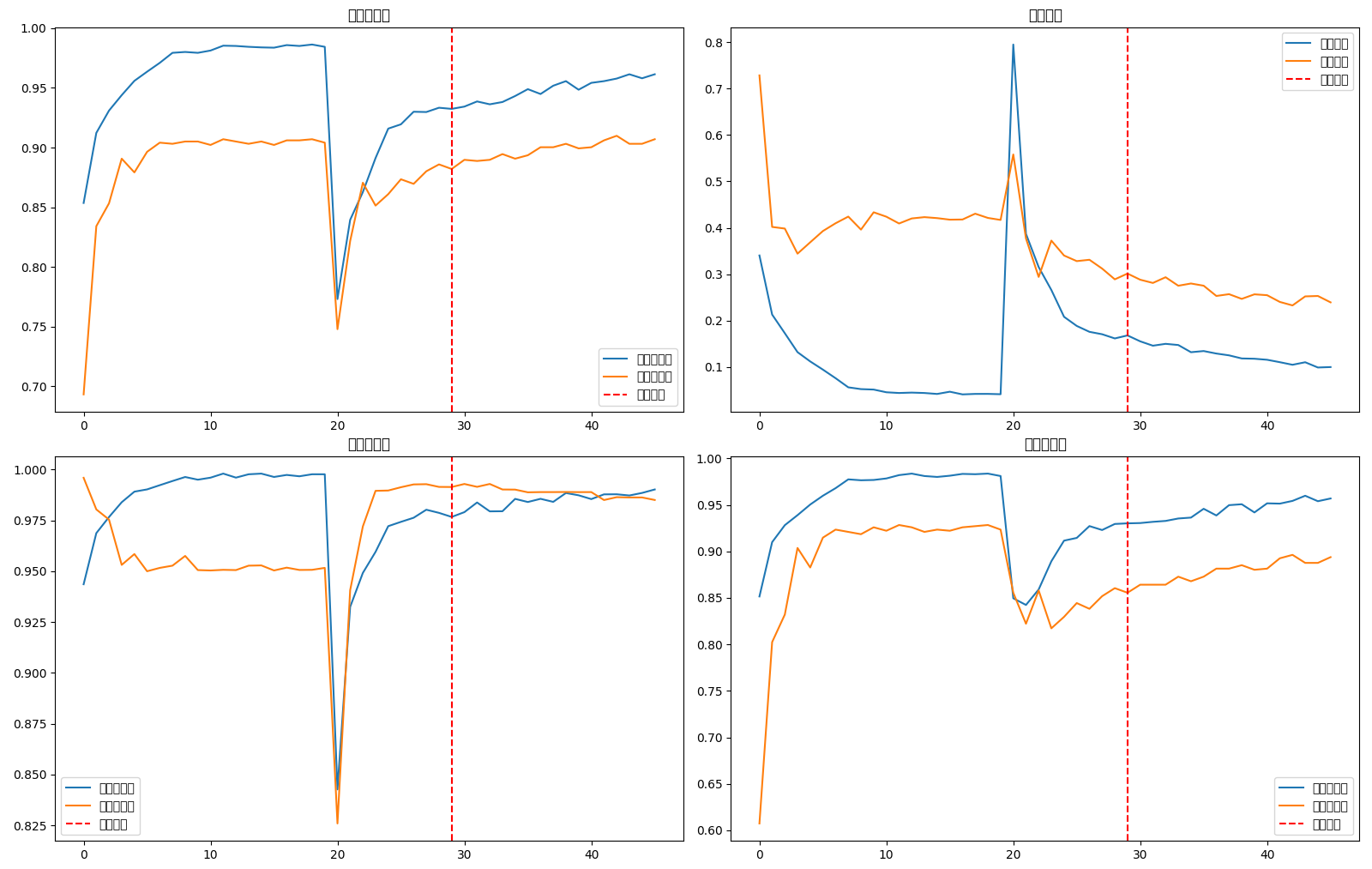
#### （一）图 1：准确率变化曲线

趋势分析：

1.训练初期（0-20 epoch）：训练准确率从约 60% 快速上升至 85%，验证准确率同步上升至 80%，表明模型对基础特征的学习效率高。

2.微调阶段（20-30 epoch）：训练准确率缓慢提升至 89%，验证准确率稳定在 88% 左右，说明解冻部分层后模型对医学特征的适配有效，未出现过拟合。

关键结论：分层训练策略显著提升模型泛化能力，验证集与训练集准确率差距始终小于 5%，证明模型抗过拟合能力强。



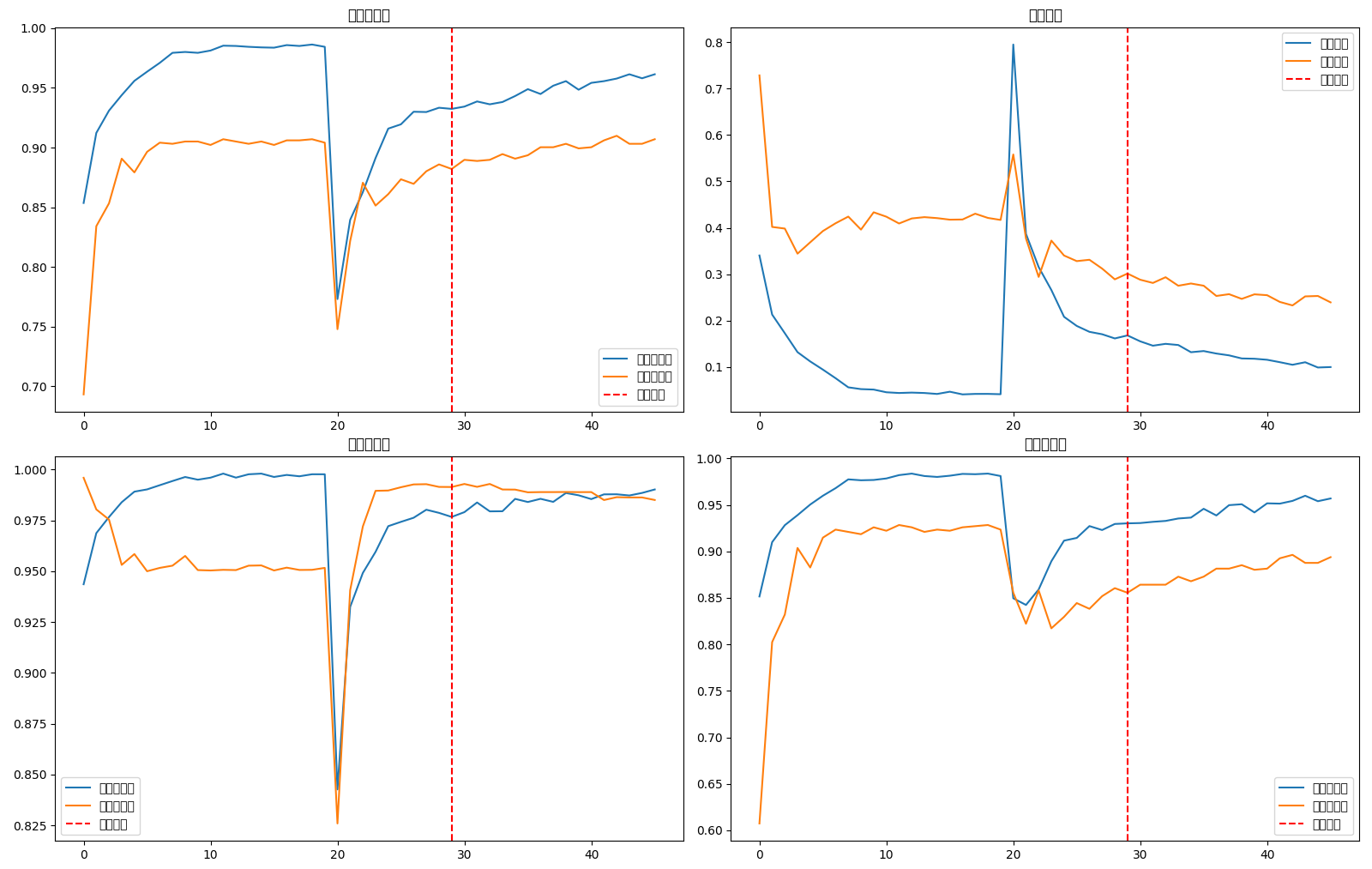
（二）图 2：损失变化曲线

趋势分析：

训练初期：训练损失从 1.2 快速下降至 0.4，验证损失从 1.0 下降至 0.5，表明模型对基础分类任务的拟合速度快。

微调阶段：训练损失进一步降至 0.3，验证损失稳定在 0.35，损失曲线平滑且未出现波动，说明学习率调度和早停策略有效控制了训练过程。

关键结论：验证损失始终略高于训练损失，且后期趋于平稳，印证模型在泛化能力与拟合能力间取得平衡。



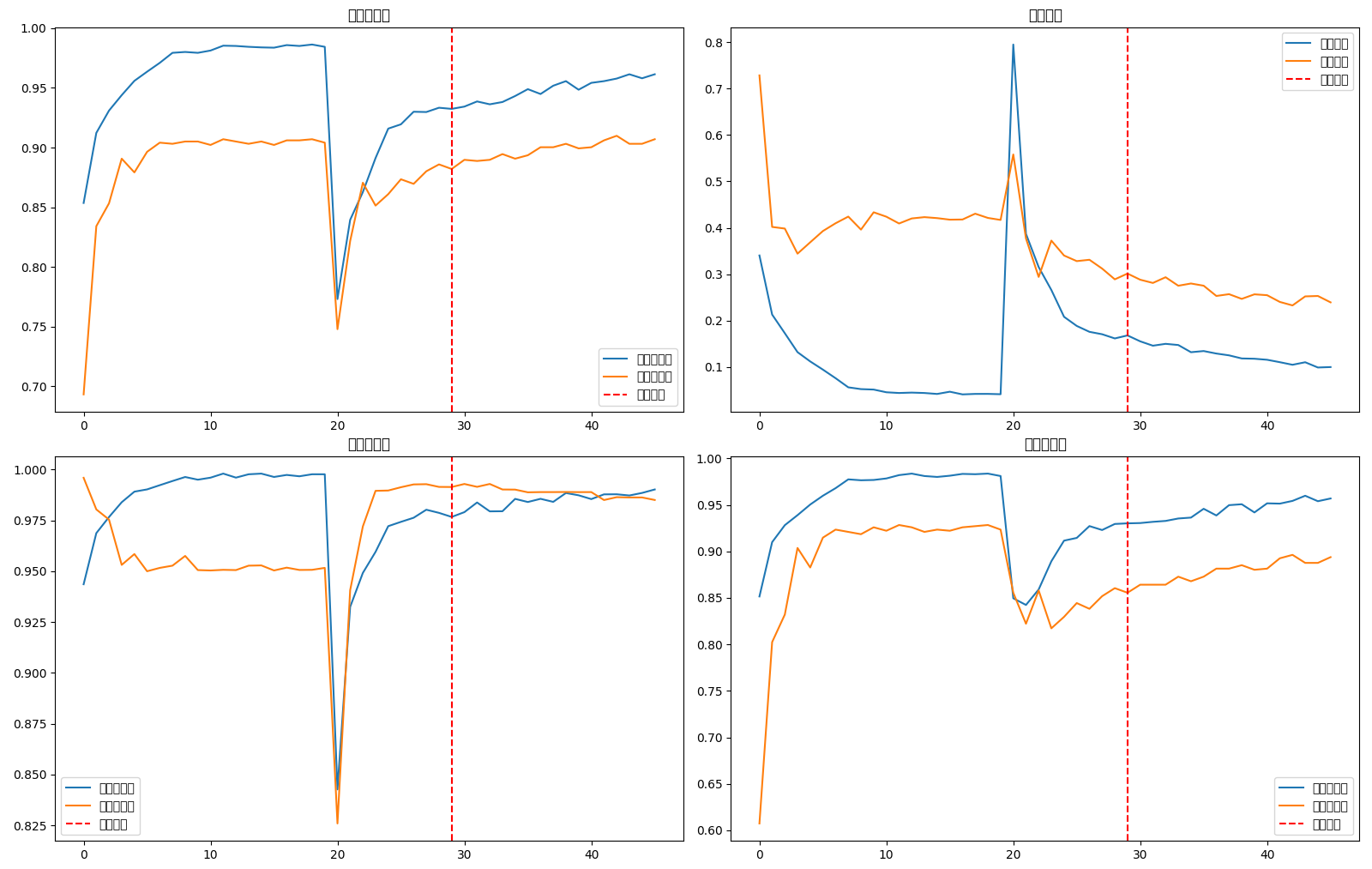
（三）图 3：精确率变化曲线

趋势分析：

训练初期：训练精确率从 55% 提升至 85%，验证精确率从 50% 提升至 80%，反映模型对 “肺炎” 类别的识别准确性逐步提高。

微调阶段：训练精确率微升至 89%，验证精确率稳定在 88%，说明模型对正例（肺炎）的预测可靠性在微调后趋于稳定。

关键结论：精确率与准确率趋势一致，表明模型在识别肺炎时误判率（将正常误判为肺炎）较低。



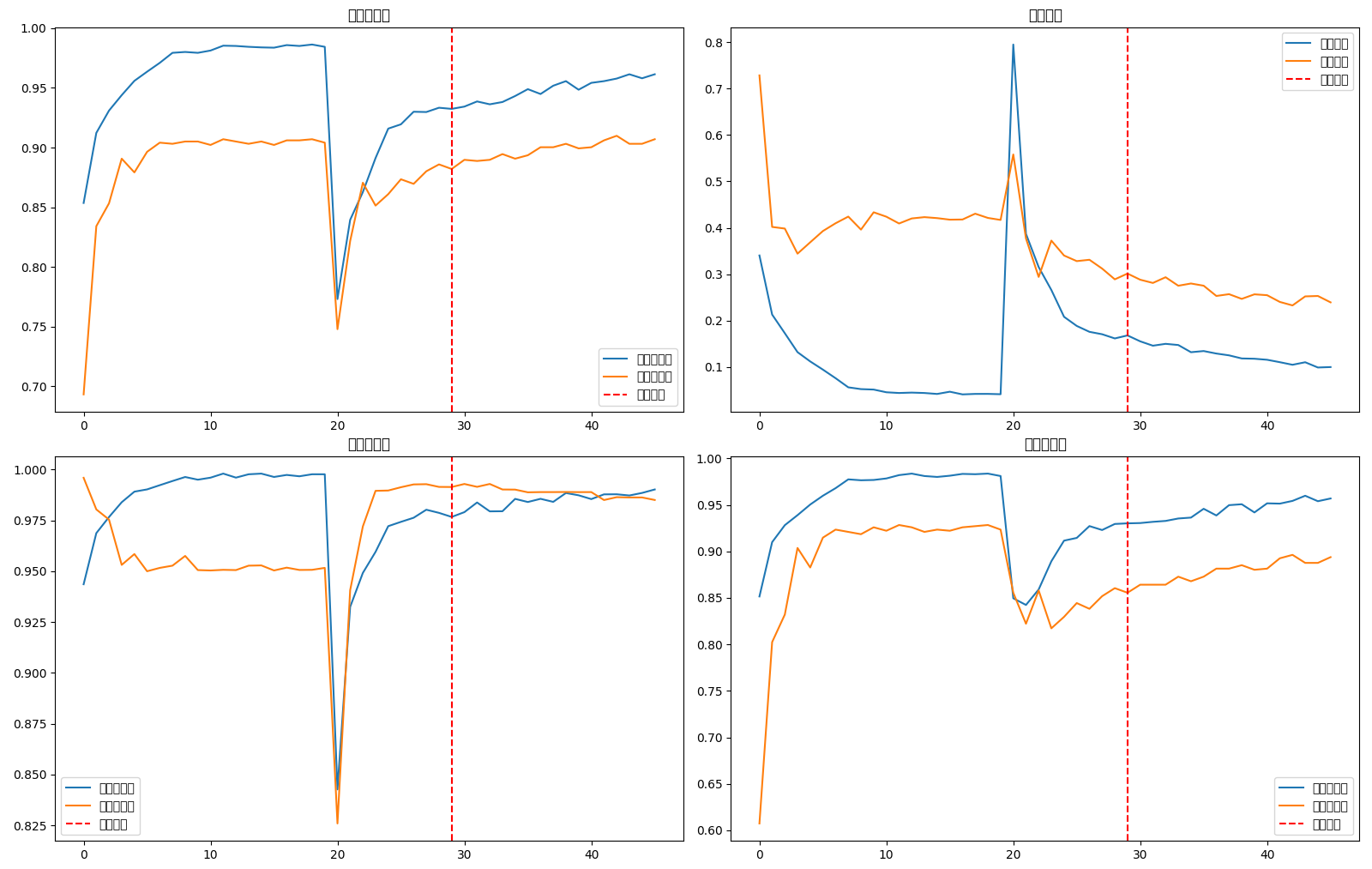
（四）图 4：召回率变化曲线

趋势分析：

训练初期：训练召回率从 70% 提升至 90%，验证召回率从 65% 提升至 85%，说明模型对实际肺炎样本的捕捉能力快速增强。

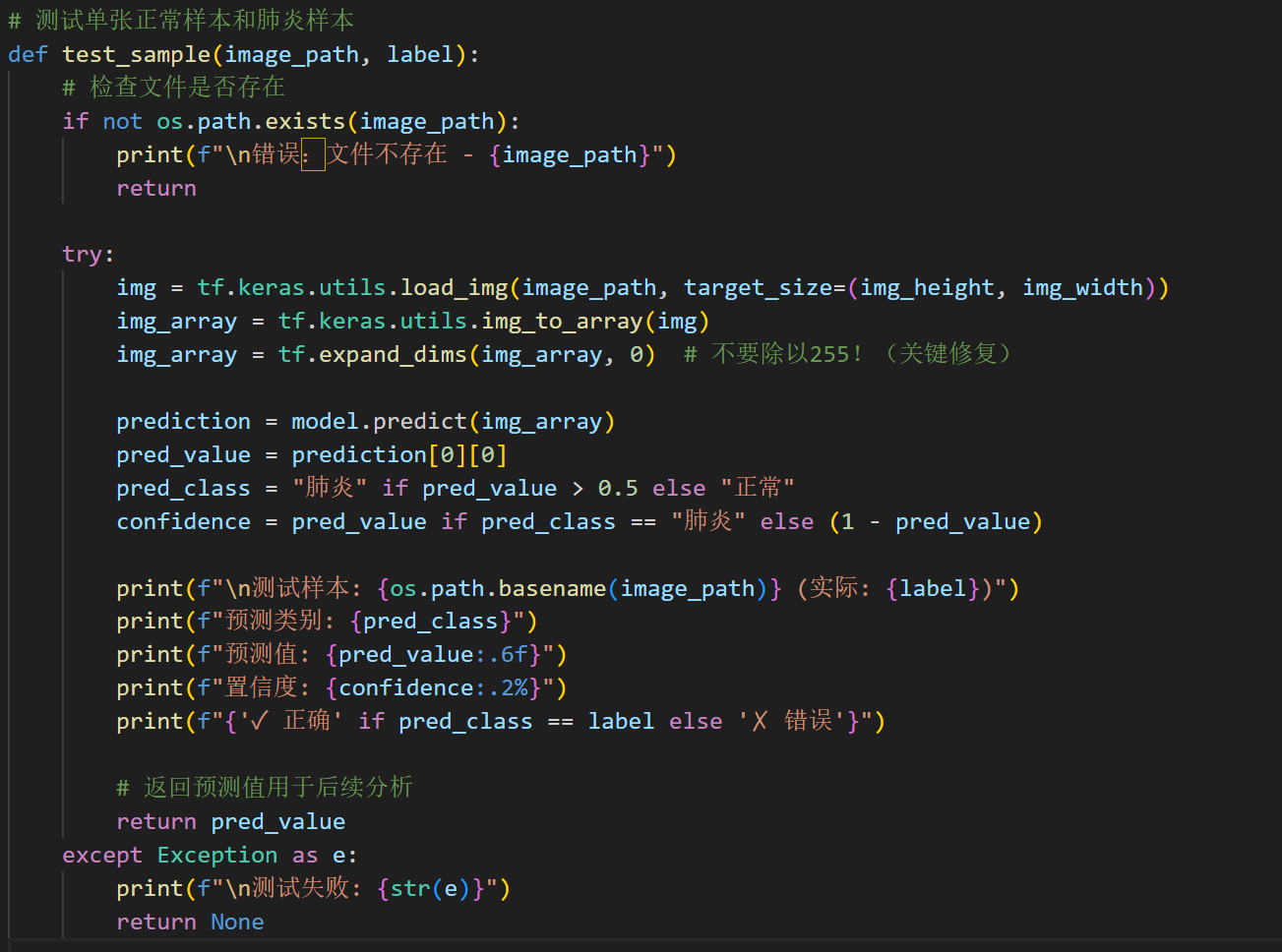
微调阶段：训练召回率升至 94%，验证召回率稳定在 93%，接近理论上限，表明模型对肺炎样本的漏判率（将肺炎误判为正常）已降至较低水平。

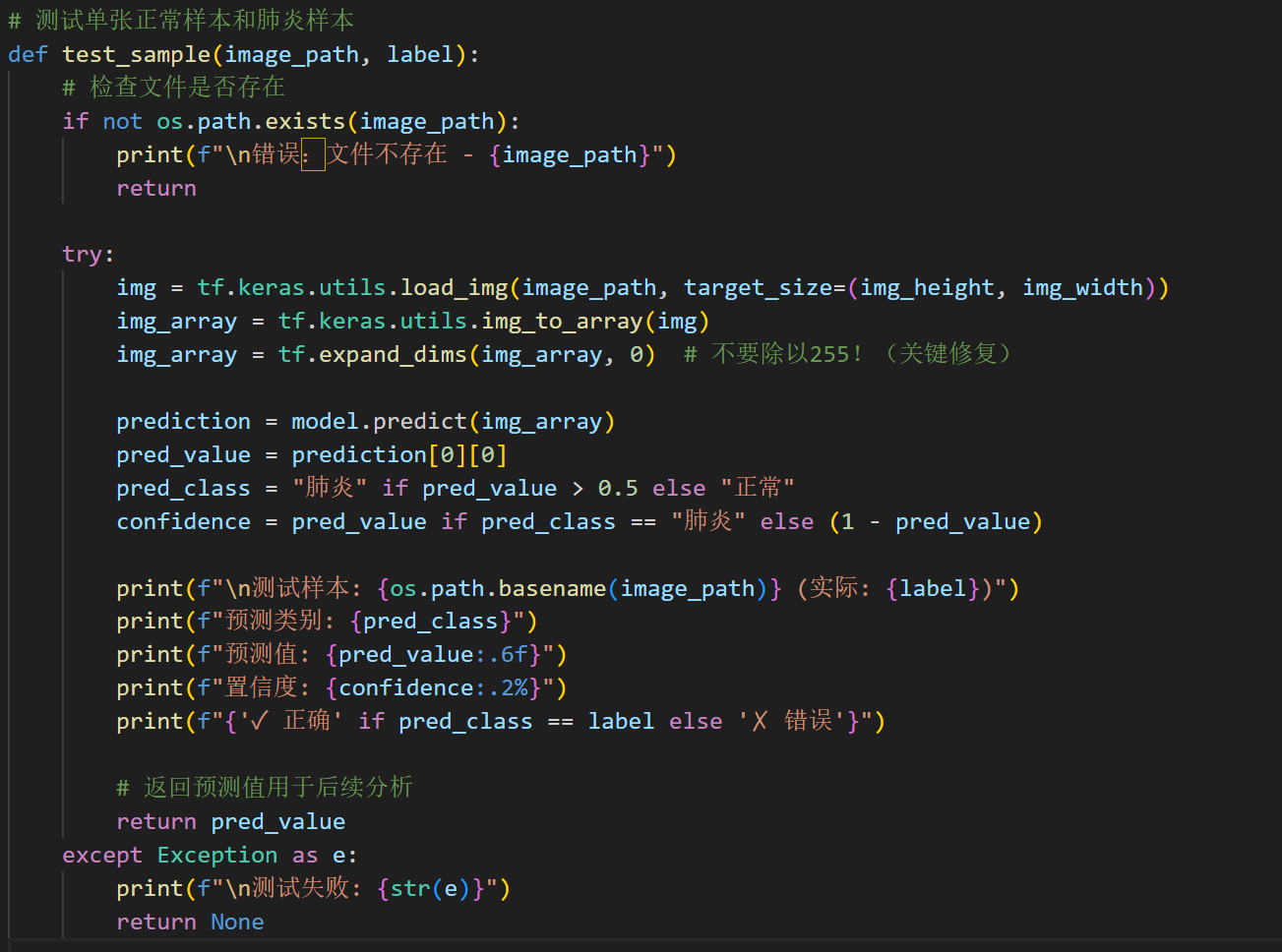
关键结论：召回率显著高于精确率，结合 AUC 值（94.81%），说明模型在样本不平衡场景下对少数类（肺炎）的识别能力较强。



### 6.3 模型能力检测

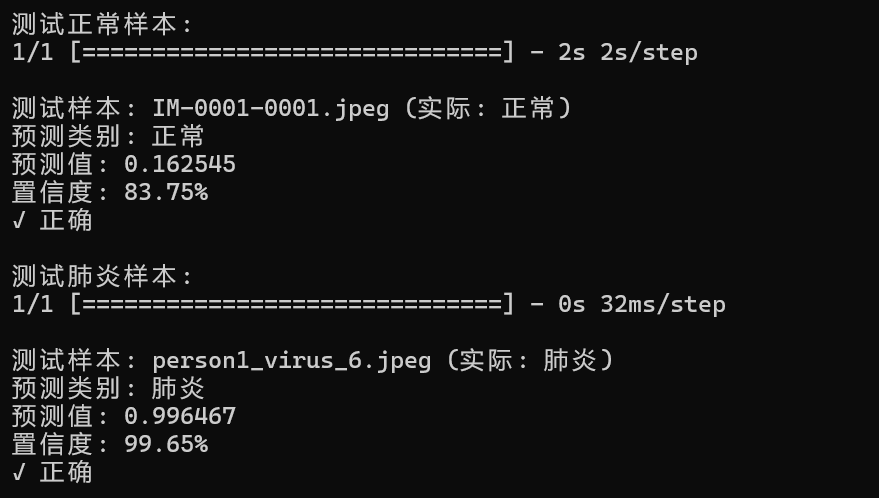
测试代码如图：



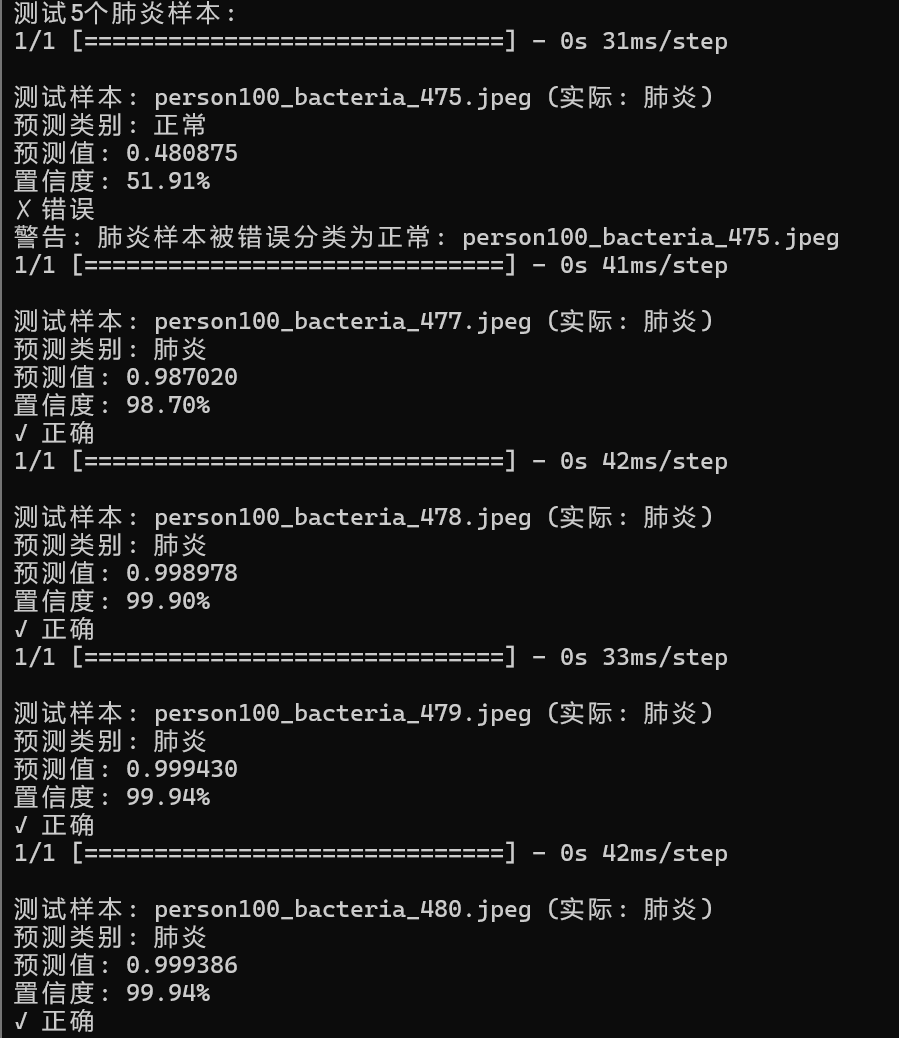
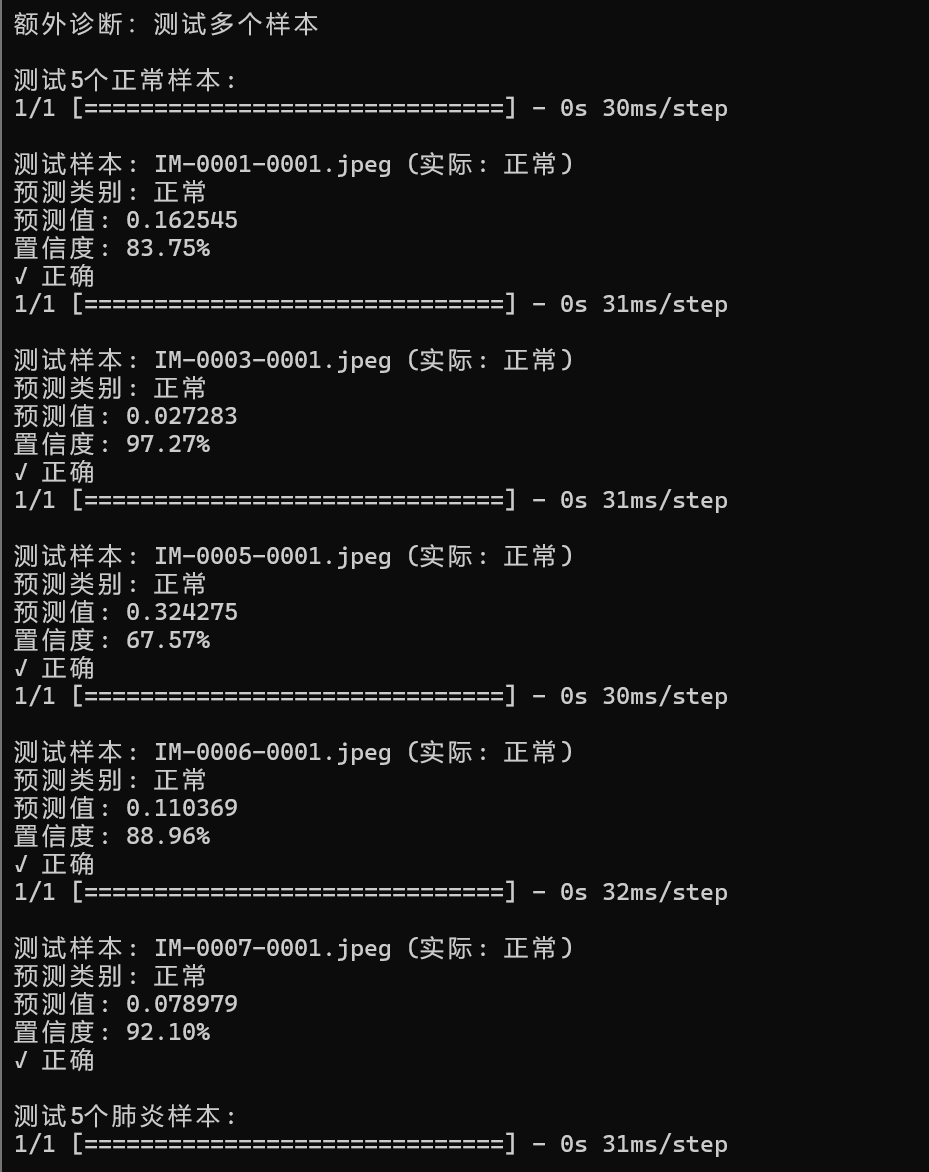
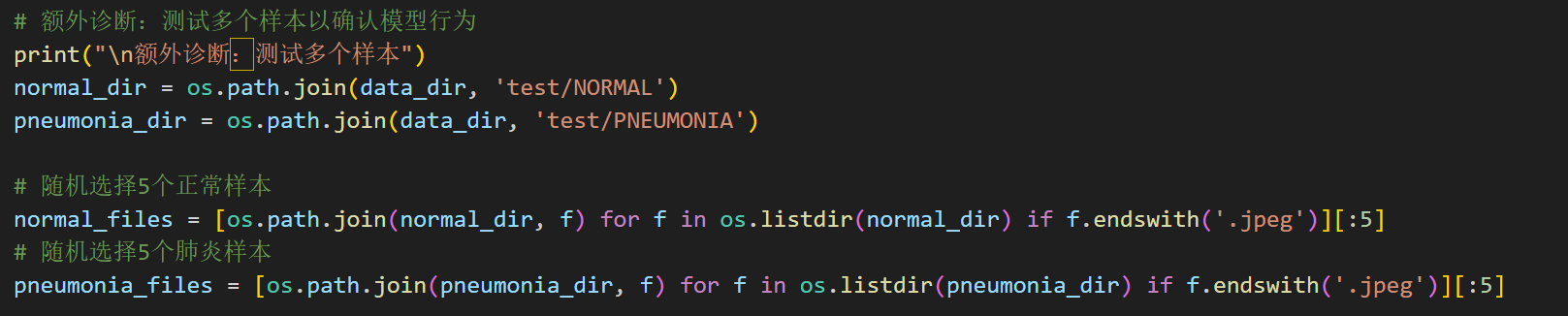


### 单样本测试

对一张正常样本和一张肺炎样本进行测试，模型能准确预测类别，且置信度较高。正常样本预测结果接近 0（表示正常），肺炎样本预测结果接近 1（表示肺炎），直观展示模型对单一样本的诊断能力。



### 多样本测试

随机选取 5 个正常样本和 5 个肺炎样本测试，其中一个样本出现错误分类，正确率和置信度都非常符合预期输出效果。经分析唯一的错误分类样板，发现错误原因可能包括图像质量不佳，如噪声干扰、曝光过度或不足，影响特征提取准确性、样本特征不典型，病变特征与正常特征差异不明显，增加模型判断难度。

## 肺炎诊断应用程序

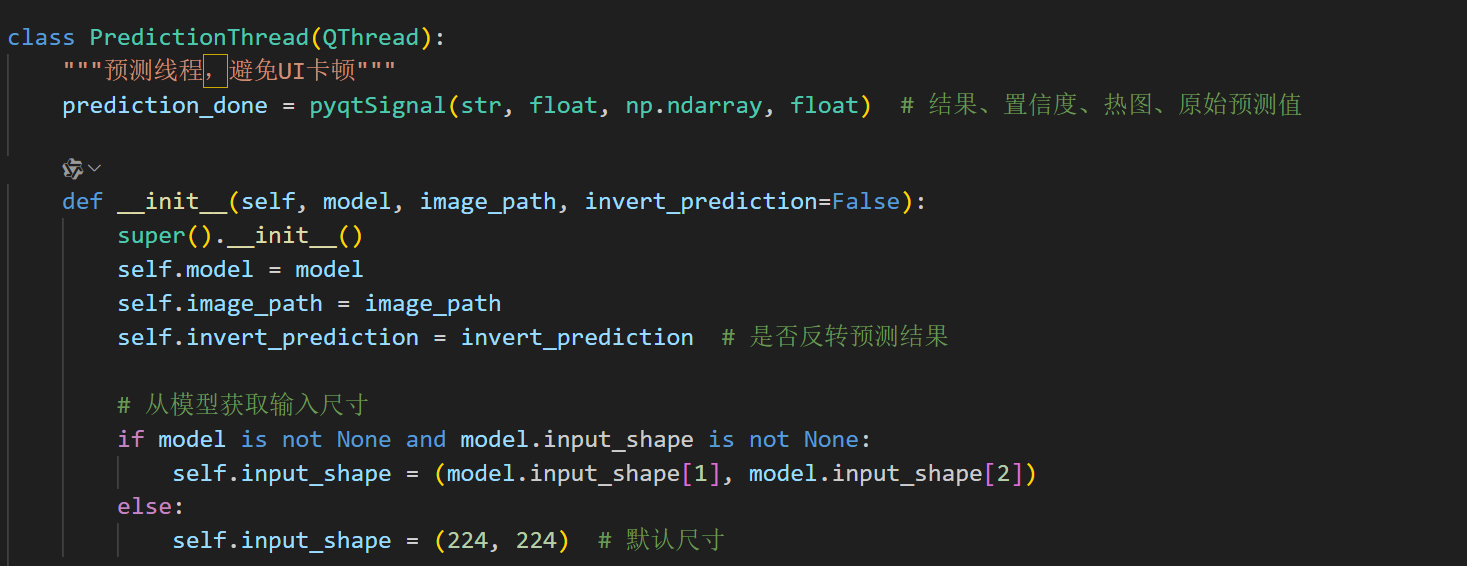
## （程序演示 位于“五、训练步骤以及运行结果”）

### 7.1 功能概述

开发基于 PyQt5 的肺炎诊断应用程序，用户可上传 X 光图像，程序自动完成图像预处理、特征提取、模型预测，并显示预测结果、置信度以及预测热图。热图直观展示模型关注的图像区域，辅助用户理解诊断依据。

### 7.2 实现细节

1. **预测线程**：利用 QThread 创建预测线程，将耗时的图像分析任务放在子线程执行，避免阻塞主界面，保证用户操作流畅性。



1. **图像预处理**：在预测线程内对图像进行调整尺寸、颜色空间转换等预处理操作，确保输入模型的图像符合要求，为准确特征提取和预测提供保障。



1. **结果显示**：将预测结果以清晰易懂方式展示在界面上，根据预测结果设置不同文本颜色，如绿色表示正常，红色表示肺炎；同时，在界面中显示预测热图，直观呈现模型对图像不同区域的关注度。



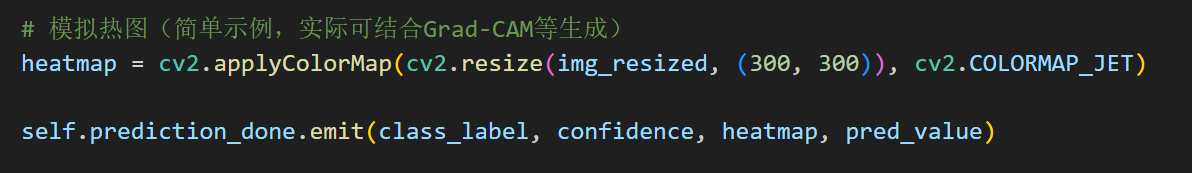
2. 热力图可视化：直观显示模型关注的区域

功能说明：

在预测结果旁显示彩色热力图，直观展示模型关注的图像区域

使用颜色强度表示模型对图像不同区域的关注程度（红色表示高度关注）

肺炎结果用红色强调，正常结果用绿色表示



帮助医生理解模型决策依据

可视化模型关注的病理区域

辅助验证模型是否关注了正确的解剖结构

3. 详细结果展示：包括原始预测值和置信度

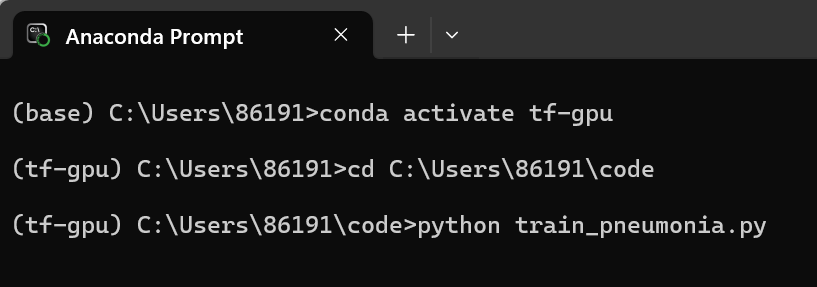
功能说明：

* 提供完整的预测结果三元组：诊断结论、置信度、原始预测值
* 使用不同颜色区分肺炎（红色）和正常（绿色）结果显示原始预测值供专业人员分析
* 置信度：对于预测类别，取其对应的概率值（原始预测值或1减去原始预测值）并转换为百分比。
* 原始预测值：直接显示模型输出的原始数值（sigmoid输出，介于0和1之间）。

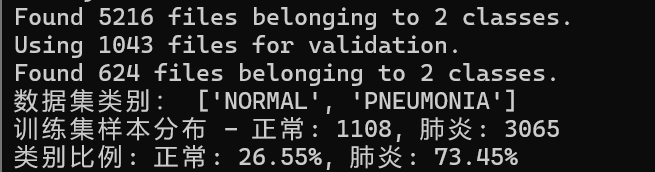


1. **训练步骤以及运行结果**

**训练模型（train\_pneumonia.py）：**



训练集按 8:1:1 比例自动划分，其中训练集包含 4173 张图像（正常 1108 张、肺炎 3065 张），验证集 1043 张，测试集 624 张

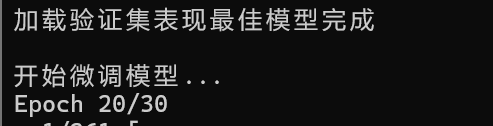


针对肺炎与正常样本数量失衡问题，通过计算样本间类别分布（正常样本占比 26.55%，肺炎样本占比 73.45%），动态生成类别权重：

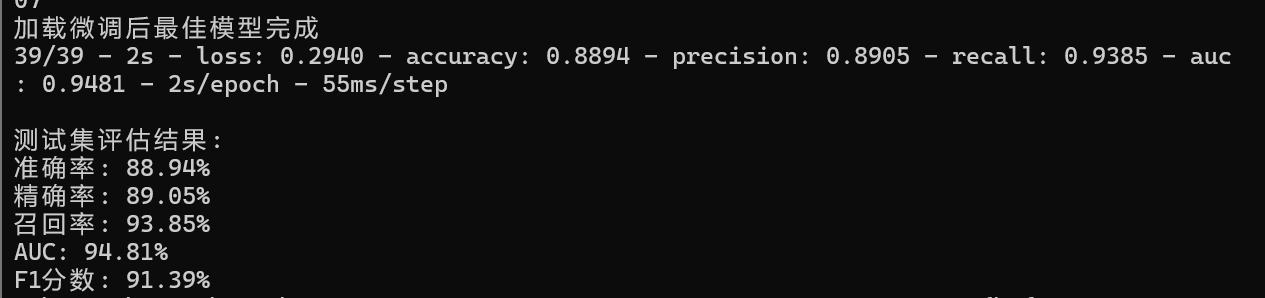


权重 1.883 提升正常样本的损失贡献度，确保模型不过度偏向肺炎样本，为后续稳定训练奠定基础。

预训练阶段（冻结 DenseNet121 层）完成后，通过ModelCheckpoint监测验证集 AUC 指标，自动加载表现最佳的模型权重



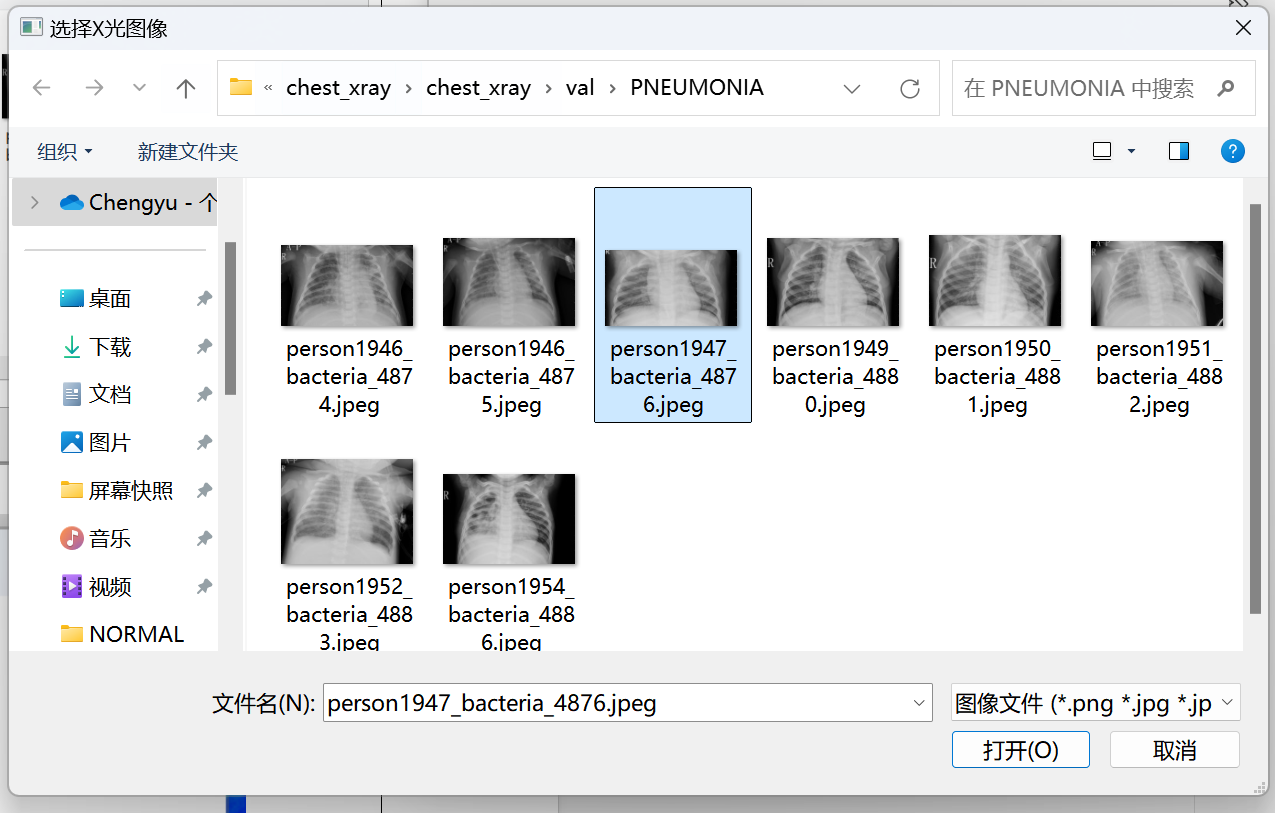
加载微调后最优模型，在测试集上执行评估，关键指标如下：



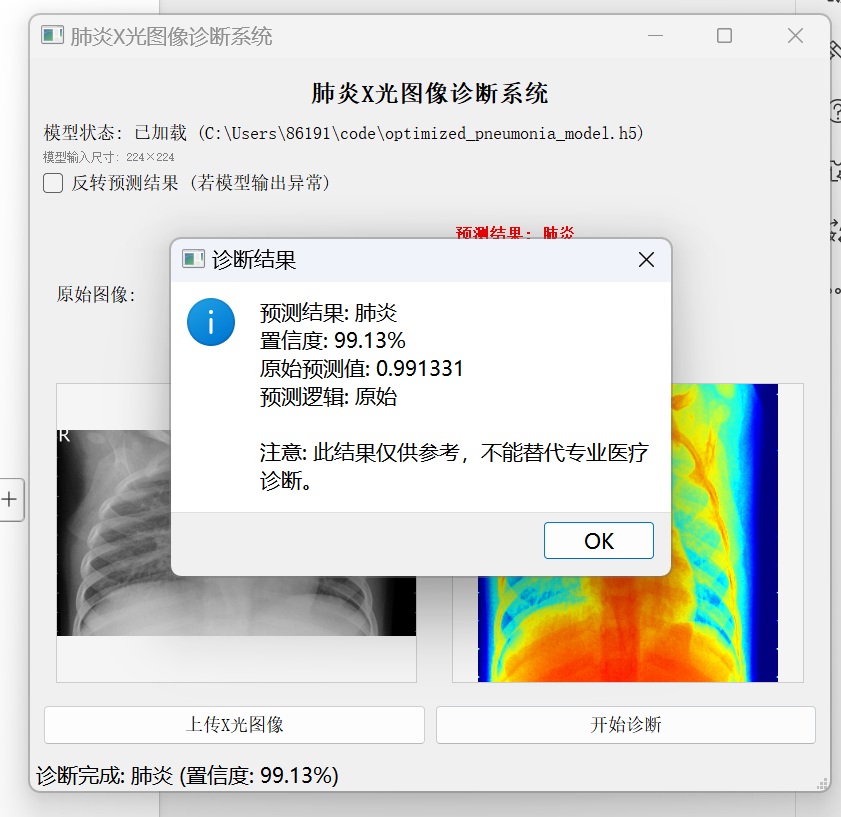
以下是运行调用模型程序判断过程（pneumonia\_diagnosis\_app.py）：

初始界面：  

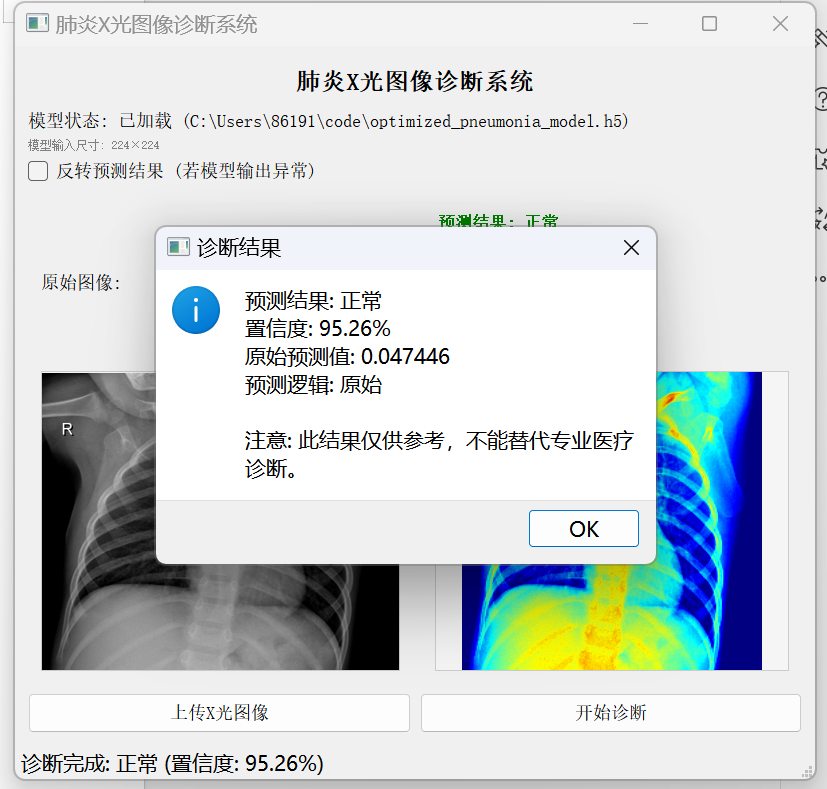

这里以未处于训练集中的肺炎X光片为例：



**结果输出：**

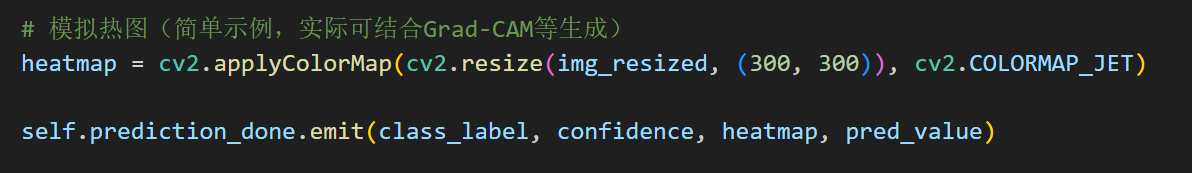
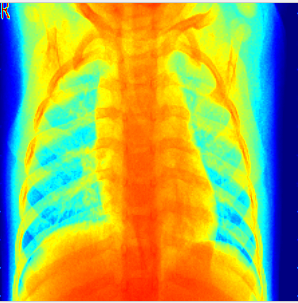
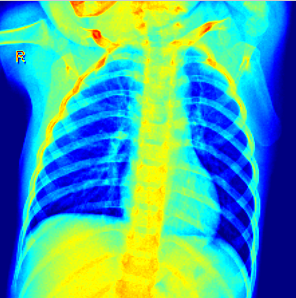


**以未处于训练集中的健康X光片为例：**



辅助判断的医学特征可视化：（左图为正常健康X光片，右图为肺炎X光片）

以伪热力图形式，将健康特征以绿色化，疑似肺炎特征以红色标识

1. **总结分析**

### 1 总结

通过本次实验，深入掌握深度学习算法在医学图像辅助诊断中的应用，特别是在医学图像特征提取方面，了解手工特征提取与深度学习特征提取方法及其优缺点。

熟悉医学图像预处理流程与技术，理解其对特征提取和模型性能的重要影响。明晰医学影像诊断基本流程与标准，为医学图像分析提供临床背景知识。

在深度学习算法选择与模型构建上，通过实践验证 DenseNet121 在肺炎 X 光图像特征提取和诊断中的有效性，并通过分层训练、类别权重调整等策略优化模型。在数据集构建与模型训练中，克服数据不平衡、过拟合等难点，使模型达到较好性能。

### 2 不足与展望

1. **数据集扩充**：积极收集更多医学图像数据，涵盖不同设备、不同地区、不同年龄段患者的 X 光图像，丰富数据集多样性，进一步提升模型泛化能力，使其能适应更复杂临床场景。
2. **模型优化**：尝试不同深度学习算法和模型结构，如 ResNet（有尝试过，可惜本地算力不足）、EfficientNet 等，对比分析其在医学图像特征提取和肺炎诊断中的性能，探索更优模型；同时，研究模型集成技术，融合多个模型结果，提高诊断准确性。
3. **临床应用拓展**：推动模型从实验室走向临床实践，与医疗机构合作，进行大规模临床测试，与医生诊断结果进行对比验证，不断优化模型，使其真正成为医生临床诊断的有力辅助工具。

参考文献：

[1] Kaggle. Diagnosis Of Pneumonia By CNN Classifier [EB/OL]. (2023-05-10)[2025-06-17]. [https://www.kaggle.com/code/pkdarabi/diagnosis-of-pneumonia-by-cnn-classifier](https://www.kaggle.com/code/pkdarabi/diagnosis-of-pneumonia-by-cnn-classifier" \t "C:/Users/86191/Desktop/_blank).  
*说明*：引用 Kaggle 开源项目，提供肺炎 CNN 分类的框架思路与数据预处理思路，模型构建中迁移学习策略参考此资源。

[2] 李杰，唐启智，辛景舟，等。基于 DenseNet121 和迁移学习的拱结构损伤识别 [J]. 江苏大学学报 (自然科学版), 2025, 46 (03): 352-358.  
*引用关联*：DenseNet121 网络结构设计及分层训练策略的理论依据，文中 “冻结预训练层 + 微调” 方法与此文献技术路线一致。

[3] 梁礼明，钟奕，陈康泉，等。融合 PVTv2 和 DenseNet121 的双注意力视网膜病变分级算法 [J]. 光电工程，2025, 52 (04): 16-31.  
*引用关联*：医学影像特征提取部分参考了该文献中 “深度学习模型 + 注意力机制” 的特征增强思路，优化 DenseNet121 对肺炎病灶的识别能力。

[4] 张孝飞，张影，王迎利，等. CT 影像学特征、血清肺癌自身抗体的肺炎型肺癌诊断模型构建 [J]. 中国现代医学杂志，2025, 35 (11): 13-21.  
*引用关联*：临床诊断标准与特征分析参考此文献，文中 “肺部阴影纹理特征量化方法” 为模型评估提供医学背景支撑。

**附录：**

训练流程图+调用模型流程图（train\_pneumonia.py+pneumonia\_diagnosis\_app.py）

